Draft Framework for Evaluating Syndromic Surveillance Systems

Daniel M. Sosin

ABSTRACT *Interest in public health surveillance to detect outbreaks from terrorism is driving the exploration of nontraditional data sources and development of new performance priorities for surveillance systems. A draft framework for evaluating syndromic surveillance systems will help researchers and public health practitioners working on nontraditional surveillance to review their work in a systematic way and communicate their efforts. The framework will also guide public health practitioners in their efforts to compare and contrast aspects of syndromic surveillance systems and decide whether and how to develop and maintain such systems. In addition, a common framework will allow the identification and prioritization of research and evaluation needs. The evaluation framework is comprised of five components: a thorough description of the system (e.g., purpose, stakeholders, how the system works); system performance experience (e.g., usefulness, acceptability to stakeholders, generalizability to other settings, operating stability, costs); capacity for outbreak detection (e.g., flexibility to adapt to changing risks and data inputs, sensitivity to detect outbreaks, predictive value of system alarms for true outbreaks, timeliness of detection); assessment of data quality (e.g., representativeness of the population covered by the system, completeness of data capture, reliability of data captured over time); and conclusions and recommendations. The draft framework is intended to evolve into guidance to support public health practice for terrorism preparedness and outbreak detection.*

KEYWORDS *Evaluation, Nontraditional surveillance, Syndromic surveillance.*

INTRODUCTION

Public health surveillance is the ongoing, systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice. Public health surveillance allows stakeholders to estimate the magnitude of public health problems, describe the natural history of a disease, determine the distribution and spread of illness, detect outbreaks, stimulate research, evaluate public health practice, monitor changes in disease agents, detect changes in health practices, and facilitate planning.[1] Historically, outbreak detection has not been a primary objective of routine surveillance. With increased interest in public health surveillance to detect outbreaks from terrorism, public health partners are exploring nontraditional data sources and placing new performance demands on surveillance systems. This draft framework (Table) differs from prior

Dr. Sosin is Director, Division of Public Health Surveillance and Informatics, Epidemiology Program Office, Centers for Disease Control and Prevention.

Correspondence: Daniel M. Sosin, MD, MPH, 4770 Buford Highway, NE, Mailstop K-74, Atlanta, GA 30341-3717.

TABLE.   Framework outline

A. System description
   1. Purpose: What is the system designed to accomplish?
   2. Stakeholders: Who is the system serving?
   3. Operation: How does the system work?
       a. Use by stakeholders
       b. Case definitions
       c. Process model
       d. Data model
       e. Interoperability
       f. Detection algorithm
       g. Privacy/confidentiality
       h. Communication

B. Experience
   1. Usefulness: In what ways has the system demonstrated value relevant to public health?
   2. Acceptability: Have stakeholders been willing to contribute to and use the system?
   3. Generalizability: How readily can the system be duplicated in another location?
   4. Stability: How consistent has the system been in providing access to reproducible results?
   5. Cost: What are the resource requirements to deploy and maintain the system?

C. Outbreak detection
   1. Flexibility: How adaptable is the system to changing needs and risk thresholds?
   2. Sensitivity and predictive value positive: What proportion of true cases and outbreaks are detected by the system? What proportion of alarms triggered by the system are desired alarms (true positives)?
   3. Timeliness: How early in the disease process or outbreak is the event detected?

D. Data quality
   1. Representativeness: How well does the system reflect the population of interest?
   2. Completeness: What percentage of data is present for each record?
   3. Reliability: Are data captured consistently across the system and over time?

E. Conclusions and recommendations for use and improvement of the syndromic surveillance system

The full draft document (www.cdc.gov/epo/dphsi/phs/syndromic.htm) gives a detailed description of the evaluation framework, which is only highlighted in this document.

guidelines for evaluating public health surveillance systems because of the explicit emphasis on outbreak detection.

There are many ways to approach the new priority of public health surveillance for outbreak detection. Information technology tools and architecture, such as specified in the National Electronic Disease Surveillance System (NEDSS), can enhance traditional surveillance resources to produce more timely and sensitive outbreak detection (www.cdc.gov/nedss/). Facilitating near-real-time sharing of data among health care providers, laboratories, and public health practitioners can increase the timely use of diagnostic data to support case detection and initiate follow-up investigations.

Nontraditional data systems for real-time or near-real-time surveillance (i.e., syndromic surveillance) provide an opportunity to supplement public health surveillance information in a way that will improve outbreak detection. Unfortunately, lacking a common understanding of what "syndromic surveillance" entails and is intended to accomplish limits the ability of public health professionals to debate,

share knowledge, and promote priority research and system development. A draft framework for evaluating syndromic surveillance systems will help researchers and public health practitioners who are working on nontraditional surveillance to review their work in a systematic way and to communicate their efforts. The framework will also guide many public health practitioners to compare and contrast different syndromic surveillance models and weigh the potential role of syndromic surveillance in a larger context of public health surveillance and information needs for public health programs. Because the value of syndromic surveillance remains unproven, it may not be appropriate for all jurisdictions. Finally, a common framework will allow the identification and prioritization of research and evaluation needs.

TERMINOLOGY

Although its use of the medical term *syndrome* is imprecise, the public health term *syndromic surveillance* has been applied to systematic and ongoing collection, analysis, and interpretation of data that precede diagnosis (e.g., laboratory test requests, emergency department chief complaint, ambulance response logs, prescription drug purchases, school or work absenteeism, as well as signs and symptoms recorded during acute care visits) and that can signal a sufficient probability of an outbreak to warrant public health investigation. Syndromic surveillance data can arise from new data collection, often at the point of medical care, and from existing data streams that are monitored for indicative disease patterns. Related terms include prediagnosis surveillance, nontraditional surveillance, enhanced surveillance, drop-in surveillance,* health indicator surveillance, and disease early warning systems.

Syndromic surveillance is already used for two conditions that illustrate the terminology. Acute flaccid paralysis is a marker of poliomyelitis and is used to detect potential cases of polio in a timely way so that each case can be investigated and confirmed without first waiting for laboratory confirmation. Here, a relatively uncommon and serious syndrome serves as a proxy for polio.[2] In contrast, influenza-like illness is a common and nonspecific marker of influenza and is tracked in sentinel settings to detect the onset of a highly predictable influenza season by watching for a rise in syndrome cases above the baseline incidence. In this instance, the system is designed to identify an aberrant pattern of disease rather than individual cases.[3]

Syndromic surveillance is only one component of surveillance preparedness for terrorism. Health care provider and laboratory technician outreach and education will enhance passive reporting of unusual diseases consistent with agents of terrorism. Furthermore, electronic laboratory reporting (i.e., the automated transfer of designated data from a laboratory database to a public health data repository) offers promise for enhancing the passive reporting of notifiable agents that might otherwise go unreported or show significant delays in reporting if not automated at the point of testing. A complete surveillance effort must support the timely inves-

---

*Drop-in surveillance is designed to be time limited in any geographic location and might be doubted as true surveillance; however, the function of monitoring high-risk events through special syndromic data collection is intended to be ongoing, only varying in location, and is therefore included here as a form of syndromic surveillance.

tigation, tracking, and provision of data that will allow appropriate management of the public health response to a terrorist event.

FRAMEWORK

This draft evaluation framework comprises five sections (Table): a thorough description of the system; system performance experience; capacity for outbreak detection; assessment of data quality; and conclusions and recommendations. An intentional overlap among sections allows individual pieces of the framework to function independently while ensuring that critical attributes and relationships are addressed. The full framework is available at http://www.cdc.gov/epo/dphsi/phs/syndromic.htm.

The system description, like a project plan, contains many categories of information and includes a statement of system purpose, clarification of the intended stakeholders, and information about how the system operates. The specific purpose(s) of the system should be acknowledged. The systems may be intended for one or more of the following purposes: (1) capturing the first or one of the early cases of disease caused by terrorism, (2) finding aberrant patterns of disease in the context of a widespread exposure, (3) tracking a proxy syndrome (e.g., fever and rash) during a known outbreak (e.g., smallpox) to identify possible cases and characterize the geographic and temporal spread of the outbreak, and (4) provide reassurance that evidence of terrorism has not been found. Most of the descriptors of system operation are found in the "Guidelines for Evaluating Public Health Surveillance Systems"[4]; however, this framework places particular emphasis on informatics in the architecture, standards, and relationships relevant to data flow. The system description ensures that stakeholders agree about the processes and priorities of the system and offers information to those considering the merits of a similar system.

The experience section addresses the historical performance of the system. The attributes to be measured and reported are usefulness, acceptability to stakeholders, generalizability to other settings, stability of operations, and costs to deploy and maintain the system. The usefulness of a syndromic surveillance system is as difficult to measure as it is for traditional disease surveillance. Ideas for measurement are suggested in the framework, but more evaluation is needed to improve the systematic measurement of usefulness. Some points that should be considered in evaluation of system usefulness are (1) the purpose(s) of the system, (2) the system's application to control of naturally occurring outbreaks, (3) the performance of traditional surveillance for the stated purpose(s), (4) a clear definition of *reassurance* when this is considered a product of the system, and (5) the limited ability of most syndromic surveillance systems to detect small outbreaks.

System costs have been highlighted in this framework because they are a key factor in understanding the value of these experimental surveillance systems. Start-up costs (e.g., for computers, software, analytic methods development) should be distinguished from ongoing costs to run the system (e.g., staff, contractual relationships for data, software licenses). Costs that can be modified (e.g., number of reporting sources, analysis and investigation resources) should specify the intensity of the resources for the given cost.

The outbreak detection section of the evaluation framework contains some of the most important and difficult conceptual issues. The attributes in this section are flexibility, sensitivity and predictive value positive (PVP), and timeliness. Flexibility refers to the adaptability of the system to meet new data collection needs and to

respond to changing priorities for detection. In periods of high risk or high concern, the detection threshold in a flexible system can be lowered to increase the detection of outbreaks at an earlier stage, albeit with the risk of more false alarms.

Sensitivity, the proportion of outbreaks in a jurisdiction that are captured by the system, and PVP, the proportion of detected events that represent outbreaks of interest, are combined in this framework because their measurement is related, and the relationship between these attributes allows comparisons of system performance. The system must be able to detect cases of the condition of interest and recognize patterns in the data to detect outbreaks. The simplest scenario occurs when every case recognized represents a potential outbreak for further investigation. Rarely, however, will the resolve and resources be sufficient to investigate every case of a reported terrorism syndrome, so interest is high in methods that scan data with optimal sensitivity and PVP. Receiver operating characteristic (ROC) curves or similar measures should be used to compare system performance at different detection thresholds. Standardized case definitions should be used when possible to distinguish system performance from that of the case definition.

Timeliness is a central attribute for outbreak detection, and efforts to improve timeliness affect all system attributes. There is considerable variability in how timeliness is measured, and the terminology and measurement of this key attribute need to be standardized. Measures of timeliness must account for the time that is needed from data arrival at the health department until a data alarm response decision is made. Large volumes of data can be transmitted instantaneously, yet the time and effort to manage, analyze, and interpret those data increase with the volume and complexity of the data. Simply measuring how long it takes to transfer data from an external database into a health department database is insufficient for understanding the timeliness of the system.

Data quality includes the traditional surveillance evaluation attributes of representativeness, completeness, and reliability. The importance of representativeness is debatable; however, skewed samples may be relevant to early detection, beyond the loss of sensitivity in an incomplete sample. For example, a system that captures data only from an insured population may miss an outbreak occurring primarily among uninsured persons.

Ultimately, the value of evaluation comes from the conclusions and recommendations drawn from and shared by the data-driven evaluation process. Strengths and weaknesses should be listed, recommendations for how to use or improve the system should be explicit, and information gaps and research or demonstration needs should be clearly described.

DISCUSSION

This draft framework describes a range of measurements important for understanding and comparing the performance of syndromic surveillance systems. The draft framework, however, provides limited guidance on standard measurements for performance that allows for system comparisons and estimation of value. Evaluation of surveillance system performance for outbreak detection is a key area for scientific investigation. Validated and standardized components of syndromic surveillance systems are needed to enhance comparisons (e.g., case definitions, detection algorithms, analysis and investigation procedures). Another area of exploration is improving and standardizing the use of simulations and test data sets to assess how well a system will detect relevant scenarios for terrorism outbreaks.

This draft framework for evaluating syndromic surveillance systems provides a focal point for evaluating and understanding syndromic surveillance systems for terrorism preparedness and outbreak detection. The draft is a work in progress by interested public health practitioners from the Centers for Disease Control and Prevention and its partners in federal, state, and local health agencies, academia, the business sector, and the military. Ongoing discussion and debate about the framework is encouraged so that this framework of issues can evolve into practical guidance for evaluation.

ACKNOWLEDGEMENT

REFERENCES

1. Thacker SB. Historical development. In: Teutsch SM, Churchill RE, eds. *Principles and Practice of Public Health Surveillance*. 2nd ed. New York, NY: Oxford University Press; 2000:1–16.
2. Robertson SE, Suleiman AJM, Mehta FR, Al-Dhahry SHS, El-Bualy MS. Poliomyelitis in Oman: acute flaccid paralysis surveillance leading to early detection and rapid response to a type 3 outbreak. *Bull World Health Organization*. 1994;72:907–914.
3. Carrat F, Flahault A, Boussard E, Farran N, Dangoumau L, Valleron AJ. Surveillance of influenza-like illness in France. The example of the 1995/1996 epidemic. *J Epidemiol Community Health*. 1998:52(suppl 1):32S–38S.
4. Centers for Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. *MMWR Morb Mortal Wkly Rep*. 2001;50:1–35.